RAW SEQUENCE LISTING
ERROR REPORT
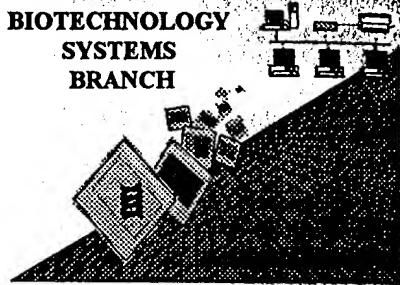
The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following CRF diskette:

Application Serial Number: _08/956,518_

Art Unit / Team No. : _1646_

Date Processed by STIC: _1/21/99_

THE ATTACHED PRINTOUT EXPLAINS THE **ERRORS DETECTED.**

PLEASE BE SURE TO FORWARD THIS INFORMATION TO THE APPLICANTS BY EITHER:

1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANTS ALONG WITH A NOTICE TO COMPLY or,

2) CALLING APPLICANTS AND FAXING THEM A COPY OF THE PRINTOUT WITH A NOTICE TO COMPLY

THIS WILL INSURE THAT THE NEXT SUBMISSION RECEIVED FROM THEM WILL BE ERROR FREE.

IF YOU HAVE ANY FURTHER QUESTIONS, PLEASE CALL:

ARTI SHAH 703-308-4212

# Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 08/956,578

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 ___ Wrapped Nucleics
The number/text at the end of each line "wrapped" down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

2 ___ Wrapped Aminos
The amino acid number/text at the end of each line "wrapped" down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".

3 ___ Incorrect Line Length
The rules require that a line not exceed 72 characters in length. This includes spaces.
All text must be visible on page.

4 ___ Misaligned Amino Acid
Numbering
.The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs
between the numbering. It is recommended to delete any tabs and uses spacing between the numbers.

5 ___ Non-ASCII
This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.

6 ___ Variable Length
Sequence(s) ___ contain n's or Xaa's which represented more than one residue.
As per the rules, each n or Xaa can only represent a single residue.
Please present the maximum number of each residue having variable length and
indicate in the (ix) features section that some may be missing.

7 ___ Wrong Designation
Sequence(s) ___ contain amino acid or nucleic acid designators which are not standard
representations as per the Sequence Rules (Please refer to paragraph 1.822)

8 ___ Skipped Sequences
(OLD RULES)
Sequence(s) ___ missing. If intentional, please use the following format for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X:
(I) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")
(xI) SEQUENCE DESCRIPTION:SEQ ID NO:X:
This sequence is intentionally skipped

Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).

9 ___ Skipped Sequences
(NEW RULES)
Sequence(s) ___ missing. If intentional, please use the following format for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

10 ___ Use of N's or Xaa's
(NEW RULES)
Use of N's and/or Xaa's have been detected in the Sequence Listing.
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.

11 ___ Use of <213>Organism
(NEW RULES)
Sequence(s) ___ are missing this mandatory field or its response.

12 ___ Use of <220>Feature
(NEW RULES)
Sequence(s) ___ are missing the <220>Feature and associated headings.
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32)
(Sec. 1.823 of new Sequence Rules)

13 ___ Wrong Format
File submitted was in the alphabetical heading format of the Old Sequence Rules.This is invalid since the
"Requirements for Patent Applications Containing Nucleotide Sequence and/or Amino Acid Disclosures"
Federal Register Notice, Vol. 63, No. 104, June 1, 1998. p. 29620
applies to applications filed on or after July 1, 1998.
AKS-Biotechnology Systems Branch- 7/10/98
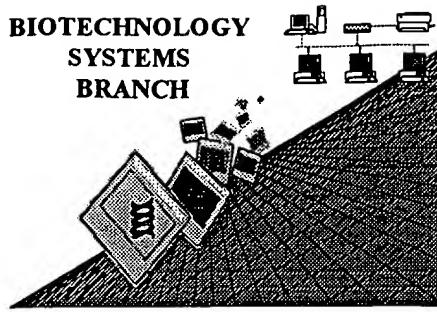
Notice of Availability of Checker Program

Applicant Aid for Biotechnology Computer Readable Form (CRF)
Sequence Listing Submissions

The Patent and Trademark Office (PTO) has developed a computer program, called Checker, that will aid applicants in identifying and correcting errors prior to making submissions for compliance with the Requirements for Patent Applications Containing Nucleotide Sequence and/or Amino Acid Sequence Disclosures (Sequence Rules: 37CRF 1.821 through 1.825).
Final rules were published in the *Federal Register* (55 FR18230) on May 1, 1990, and in the PTO *Official Gazette* (1114 Off.Gaz.PatOffice 29) on May 15, 1990.

Checker is a DOS-based software program that is intended to assist users in determining whether errors may be present in the sequence listings, and is not intended to guarantee that the submission is error-free.

The most current version of the software is available via computer downloading, details are below. Copies on diskette are also available. Updated software versions will not be automatically mailed out; any updates will be announced in the PTO *Official Gazette*.

The software can be accessed/requested from the following locations:

1) Dial-up access through the Internet. Location is ftp://ftp.uspto.gov
   The software is in current directory: pub/checker/
   Download all the files. Cost: Free-of-charge

3) For diskette copies, mail to: U.S.P.T.O., OEIP, CRYSTAL PARK 3, SUITE 441
                                  WASHINGTON DC 20231

   COST FOR DISKETTE IS $ 25.00
   METHOD OF PAYMENT:
       Check payable to Commissioner of Patents and Trademarks
       VISA/ Mastercard/ Charge- Charges can be faxed to 703-306-2737
       PTO Deposit Account

For Further Information, Contact: **Arti Shah at 703-308-4212**

> **This Raw Listing contains the General Information Section and those Sequences containing ERRORS.**

Corrected Diskette Needed / Does Not Comply

```
 1                              SEQUENCE LISTING
 2
 3   (1)     General Information:
 4
 5       (i) APPLICANT: Leonard, Sherry
 6                      Freedman, Robert
 7
 8      (ii) TITLE OF INVENTION: ALPHA-7 NICOTINIC RECEPTOR
 9
10     (iii) NUMBER OF SEQUENCES: 121
11
12      (iv) CORRESPONDENCE ADDRESS:
13           (A) ADDRESSEE: MEDLEN & CARROLL, LLP
14           (B) STREET: 220 Montgomery Street, Suite 2200
15           (C) CITY: San Francisco
16           (D) STATE: CA
17           (E) COUNTRY: USA
18           (F) ZIP: 94104
19
20       (v) COMPUTER READABLE FORM:
21           (A) MEDIUM TYPE: Floppy disk
22           (B) COMPUTER: IBM PC compatible
23           (C) OPERATING SYSTEM: PC-DOS/MS-DOS
24           (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
25
26      (vi) CURRENT APPLICATION DATA:
27           (A) APPLICATION NUMBER: US 08/956,518
28           (B) FILING DATE: 23-OCT-1997
29           (C) CLASSIFICATION:
30
31   (viii) ATTORNEY/AGENT INFORMATION:
32           (A) NAME: MacKnight, Kamrin T.
33           (B) REGISTRATION NUMBER: 38,230
34           (C) REFERENCE/DOCKET NUMBER: UTC-03042
35
36      (ix) TELECOMMUNICATION INFORMATION:
37           (A) TELEPHONE: 415-705-8410
38           (B) TELEFAX: 415-397-8338
39
40
```

ERRORED SEQUENCES FOLLOW:

INPUT SET: S30866.raw

```
185    (2) INFORMATION FOR SEQ ID NO:9:
186
187         (i) SEQUENCE CHARACTERISTICS:
188             (A) LENGTH: 337 base pairs
189             (B) TYPE: nucleic acid
190             (C) STRANDEDNESS: single
191             (D) TOPOLOGY: linear
192
193        (ii) MOLECULE TYPE: other nucleic acid
194             (A) DESCRIPTION: /desc = "DNA"
195
196
197
198
199        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:
200
201    AGATGCCCAA GTGGACCAGA GTCATCCTTC TGAACTGGTG CGCGTGGTTC
202    CTGCGAATGA       60
203
204    AGAGGCCCGG GGAGGACAAG GTGCGCCCGG CCTGCCAGCA CAAGCAGCGG
205    CGCTGCAGCC      120
206
207    TGGCCAGTGT GGAGATGAGC GCCGTGGGCC CGCCGCCCGC CAGCAACGGG
208    AACCTGCTGT      180
209
210    ACATCGGCTT CCGCGGCCTG GACGGCGTGC ACTGTGTCCC GACCCCCGAC
211    TCTGGGGTAG      240
212
213    TGTGTGGCCG CATGGCCTGC TCCCCCACGC ACGATGAGCA CCTCCTGCAC
214    GGCGGGCAAC      300
215
216    CCCCCGAGGG GGACCCGGAC TTGGCCAAGA TCCTGGA                     337
217
```

```
236    (2) INFORMATION FOR SEQ ID NO:11:
237
238         (i) SEQUENCE CHARACTERISTICS:
239             (A) LENGTH: 54 base pairs
240             (B) TYPE: nucleic acid
241             (C) STRANDEDNESS: single
242             (D) TOPOLOGY: linear
243
244        (ii) MOLECULE TYPE: other nucleic acid
245             (A) DESCRIPTION: /desc = "DNA"
246
247
248
249
250        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:
251
252    CACACACACA TCACACACAC ACACACACAC ACACATACAC ACACACACCA CACA
253              54
254
```

```
921    (2) INFORMATION FOR SEQ ID NO:49:
922
923        (i) SEQUENCE CHARACTERISTICS:
924            (A) LENGTH: 48 base pairs
925            (B) TYPE: nucleic acid
926            (C) STRANDEDNESS: single
927            (D) TOPOLOGY: linear
928
929        (ii) MOLECULE TYPE: other nucleic acid
930            (A) DESCRIPTION: /desc = "DNA"
931
932
933
934
935        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:
936
937    CAUCAUCAUC AUCCAGCGTA CATCGATGTA GCAGGAACTC TTGAATAT
938        48
939
```

same

```
940    (2) INFORMATION FOR SEQ ID NO:50:
941
942        (i) SEQUENCE CHARACTERISTICS:
943            (A) LENGTH: 41 base pairs
944            (B) TYPE: nucleic acid
945            (C) STRANDEDNESS: single
946            (D) TOPOLOGY: linear
947
948        (ii) MOLECULE TYPE: other nucleic acid
949            (A) DESCRIPTION: /desc = "DNA"
950
951
952        (ix) FEATURE:
953            (A) NAME/KEY: -
954            (B) LOCATION: 32..33
955            (D) OTHER INFORMATION: /note= "The residue at this
956    position is Inosine."
957
958        (ix) FEATURE:
959            (A) NAME/KEY: -
960            (B) LOCATION: 36..37
961            (D) OTHER INFORMATION: /note= "The residue at this
962    position is Inosine."
963
964        (ix) FEATURE:
965            (A) NAME/KEY: -
966            (B) LOCATION: 41..42
967            (D) OTHER INFORMATION: /note= "The residue at this
968    position is Inosine."
969
970
971        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:
```

Handwritten annotations:
"C" at location 32 can only represent itself, "G" at location 33 only represents itself. Nothing else.

If you want location 32 to be Inosine, use "N". Same with location 33.

same error

(42)? There are only 41 nucleic acids in the sequence

same error

Inosine can only be represented by "N".

INPUT SET: S30866.raw

```
972
973    CUACUACUAC UAGGCCACGC GTCGACTAGT ACGGCGGGGG G
974    41
975
```

```
1570    (2) INFORMATION FOR SEQ ID NO:84:
1571
1572           (i) SEQUENCE CHARACTERISTICS:
-->  1573              (A) LENGTH: 55 base pairs
1574              (B) TYPE: nucleic acid
1575              (C) STRANDEDNESS: single
1576              (D) TOPOLOGY: linear
1577
1578          (ii) MOLECULE TYPE: other nucleic acid
1579              (A) DESCRIPTION: /desc = "DNA"
1580
1581
1582
1583
1584         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:
1585
1586    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNCT
-->  1587    GCACG                55
1588
```

```
1751    (2) INFORMATION FOR SEQ ID NO:94:
1752
1753           (i) SEQUENCE CHARACTERISTICS:
-->  1754              (A) LENGTH: 457 base pairs
1755              (B) TYPE: nucleic acid
1756              (C) STRANDEDNESS: single
1757              (D) TOPOLOGY: linear
1758
1759          (ii) MOLECULE TYPE: other nucleic acid
1760              (A) DESCRIPTION: /desc = "DNA"
1761
1762
1763
1764
1765         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:
1766
1767    AGAACGCAAG GGAGAGGTAG AGCCTGGCCT TGGGCAGCCC CTGGCCTGGC
-->  1768    CAGAGGCGCG         60
1769
1770    AGGCCGAGAG CCCGCTCGGT GGAGACTGGG GGTGGAGGTG CCCGGAGCGT
-->  1771    ACCCAGCGCC        120
1772
1773    GGGAGTACCT CCCGCTCACA CCTCGGGCTG CAGTTCCCTG GGTGGCCGCC
-->  1774    GAGACGCTGG        180
1775
1776    CCCGGGCTGG AGGGATGGCG GGGCGGGGAC GGGGGCGGGG GCGGGGCTCG
-->  1777    TCACGTGGAG        240
1778
```

INPUT SET: S30866.raw

```
      1779   AGGCGCGCGG GGGCGGGCGG GGCGGGGGCG CGCGCCCGGC TCCTTAAAGG
-->   1780   CGCGCGAGCC       300
      1781
      1782   GAGCGGCGAG GTGCCTCTGT GGCCGCAGGC GCAGGCCCGG GCGACAGCCG
-->   1783   AGACGTGGAG       360
      1784
      1785   CGCGCCGGCT CGCTGCAGCT CCGGGACTCA ACATGCGCTG CTCGCCGGGA
-->   1786   GGCGTCTGGC       420
      1787
      1788   TGGCGCTGGC CGCGTCGCTC CTGCACGGTA AAGCCAC                       457
      1789
```

```
      1790   (2) INFORMATION FOR SEQ ID NO:95:
      1791
      1792        (i) SEQUENCE CHARACTERISTICS:
-->   1793            (A) LENGTH: 308 base pairs
      1794            (B) TYPE: nucleic acid
      1795            (C) STRANDEDNESS: single
      1796            (D) TOPOLOGY: linear
      1797
      1798       (ii) MOLECULE TYPE: other nucleic acid
      1799            (A) DESCRIPTION: /desc = "DNA"
      1800
      1801
      1802
      1803
      1804       (xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:
      1805
      1806   CAGGCCGCCA CATAGCTCCC GCCAAGTCCT CGGTGCCCCT TGCCATTTTC
-->   1807   CAGCCGCGTC        60
      1808
      1809   CCACGAGGGT CACGGCGGCG GGGAGAGGTG GAGCCGCGAG AGCTCGGCCG
-->   1810   GGGGCCCCGC       120
      1811
      1812   CTGGTGGCCG CGGCCATGAC AGCGGCTCGG GACTGGCTCC TTTTCCGCGC
-->   1813   CCCTCCCGCC       180
      1814
      1815   GGAGGTGAGG GGAAGATGTC CATGTCAGGG TTCAAGGCCA AACCGAAGTT
-->   1816   ACTGGCCCTC       240
      1817
      1818   TATCTTCCAG GAGAACCAGG AGCCACAGCC GCGGCTCACG CCCCACCGCA
-->   1819   ACATTAAGGT       300
      1820
      1821   GAGTCGCC                                                       308
      1822
```

```
      1823   (2) INFORMATION FOR SEQ ID NO:96:
      1824
      1825        (i) SEQUENCE CHARACTERISTICS:
-->   1826            (A) LENGTH: 145 base pairs
      1827            (B) TYPE: nucleic acid
      1828            (C) STRANDEDNESS: single
      1829            (D) TOPOLOGY: linear
```

INPUT SET: S30866.raw

```
     1830
     1831        (ii) MOLECULE TYPE: other nucleic acid
     1832              (A) DESCRIPTION: /desc = "DNA"
     1833
     1834
     1835
     1836
     1837        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:
     1838
     1839    CTCATTTCAG ATTACAAGTG GACACCTGAG TCAGCAGGAC CTGGAATCCC
-->  1840    AGATGAGAGA         60
     1841
     1842    GCTTATCTAC ACGACTCAGA TCTTGTTGTC ACCCCCATTA TTGACAATCC
-->  1843    AAAGGTGCAG        120
     1844
     1845    AAAGCACTCT GACAAGTGAG TTGTA                                 145
     1846
```

```
     1847    (2) INFORMATION FOR SEQ ID NO:97:
     1848
     1849        (i) SEQUENCE CHARACTERISTICS:
-->  1850              (A) LENGTH: 84 base pairs
     1851              (B) TYPE: nucleic acid
     1852              (C) STRANDEDNESS: single
     1853              (D) TOPOLOGY: linear
     1854
     1855        (ii) MOLECULE TYPE: other nucleic acid
     1856              (A) DESCRIPTION: /desc = "DNA"
     1857
     1858
     1859
     1860
     1861        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:
     1862
     1863    TTAACCACAG ATAATGAAAC AACCACCATC GGTTAAATTT GATGCAAAAA
-->  1864    TATTGCATCT         60
     1865
     1866    ACCAGCATTT TCAGGTAGGA TCAT                                   84
     1867
```

```
     1868    (2) INFORMATION FOR SEQ ID NO:98:
     1869
     1870        (i) SEQUENCE CHARACTERISTICS:
-->  1871              (A) LENGTH: 67 base pairs
     1872              (B) TYPE: nucleic acid
     1873              (C) STRANDEDNESS: single
     1874              (D) TOPOLOGY: linear
     1875
     1876        (ii) MOLECULE TYPE: other nucleic acid
     1877              (A) DESCRIPTION: /desc = "DNA"
     1878
     1879
     1880
```

```
        1881
        1882        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:
        1883
        1884     TTTATTCTAG TTCCAATTGC TAATCCAGCA TTTGTGGATA GCTGCAAACT
-->     1885     GCGATATGTA         60
        1886
        1887     AGTAACA                                                         67
        1888
```

```
        1889     (2) INFORMATION FOR SEQ ID NO:99:
        1890
        1891        (i) SEQUENCE CHARACTERISTICS:
-->     1892            (A) LENGTH: 100 base pairs
        1893            (B) TYPE: nucleic acid
        1894            (C) STRANDEDNESS: single
        1895            (D) TOPOLOGY: linear
        1896
        1897        (ii) MOLECULE TYPE: other nucleic acid
        1898            (A) DESCRIPTION: /desc = "DNA"
        1899
        1900
        1901
        1902
        1903        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:
        1904
        1905     CTGTTTCTAG TGCTGATGAG CGCTTTGACG CCACATTCCA CACTAACGTG
-->     1906     TTGGTGAATT         60
        1907
        1908     CTTCTGGGCA TTGCCAGTAC CTGCCTCCAG GTAAGCTGCA                      100
        1909
```

```
        1928     (2) INFORMATION FOR SEQ ID NO:101:
        1929
        1930        (i) SEQUENCE CHARACTERISTICS:
-->     1931            (A) LENGTH: 392 base pairs
        1932            (B) TYPE: nucleic acid
        1933            (C) STRANDEDNESS: single
        1934            (D) TOPOLOGY: linear
        1935
        1936        (ii) MOLECULE TYPE: other nucleic acid
        1937            (A) DESCRIPTION: /desc = "DNA"
        1938
        1939
        1940
        1941
        1942        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:
        1943
        1944     AGAACGCAAG GGAGAGGTAG AGCCTGGCCT TGGGCAGCCC CTGGCCTGGC
-->     1945     CAGAGGCGCG         60
        1946
        1947     AGGCCGAGAG CCCGCTCGGT GGAGACTGGG GGTGGAGGTG CCCGGAGCGT
-->     1948     ACCCAGCGCC        120
        1949
```

INPUT SET: S30866.raw

```
        1950   GGGAGTACCT CCCGCTCACA CCTCGGGCTG CAGTTCCCTG GGTGGCCGCC
-->     1951   GAGACGCTGG        180
        1952
        1953   CCCGGGCTGG AGGGATGGCG GGGCGGGGAC GGGGGCGGGG GCGGGGCTCG
-->     1954   TCACGTGGAG        240
        1955
        1956   AGGCGCGCGG GGGCGGGCGG GGCGGGGGCG CGCGCCCGGC TCCTTAAAGG
-->     1957   CGCGCGAGCC        300
        1958
        1959   GAGCGGCGAG GTGCCTCTGT GGCCGCAGGC GCAGGCCCGG GCGACAGCCG
-->     1960   AGACGTGGAG        360
        1961
        1962   CGCGCCGGCT CGCTGCAGCT CCGGGACTCA AC.                         392
        1963
```

```
        1964   (2) INFORMATION FOR SEQ ID NO:102:
        1965
        1966       (i) SEQUENCE CHARACTERISTICS:
-->     1967           (A) LENGTH: 689 base pairs
        1968           (B) TYPE: nucleic acid
        1969           (C) STRANDEDNESS: single
        1970           (D) TOPOLOGY: linear
        1971
        1972      (ii) MOLECULE TYPE: other nucleic acid
        1973           (A) DESCRIPTION: /desc = "DNA"
        1974
        1975
        1976
        1977
        1978      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:
        1979
        1980   AGCCCTTTCC CAGGCGGTAG CGGGGGCAGT GGTGCTGTTG CCCTTTTAAA
-->     1981   CTGCGGCTTG         60
        1982
        1983   ACGGGAGCCG CGCCTCCTGT CGGTGGAGTC GGTTATAAAG GGAGCAGCCC
-->     1984   CGCAGGCCGC        120
        1985
        1986   CACATAGCTC CCGCCAAGTC CTCGGTGCCC CTTGCCATTT TCCAGCCGCG
-->     1987   CTCCCACGAG        180
        1988
        1989   GGTCACGGCG GCGGGGAGAG GTGGAGCCGC GAGAGCTCGG CCGGGGGCCC
-->     1990   CGCCTGGTGG        240
        1991
        1992   CCGCGGCCAT GACAGCGGCT CGGGACTGGC TCCTTTTCCG CGCCCCTCCC
-->     1993   GCCGGAGGTG        300
        1994
        1995   AGGGGAAGAT GTCCATGTCA GGGTTCAAGG CCAAACCGAA GTTACTGGCC
-->     1996   TCTATCTTCC        360
        1997
        1998   AGGAGAACCA GGAGCCACAG CCGCGGCTCA CGCCCCACCG CAACATTAAG
-->     1999   ATTACAAGTG        420
        2000
        2001   GACACCTGAG TCAGCAGGAC CTGGAATCCC AGATGAGAGA GCTTATCTAC
```

INPUT SET: S30866.raw

```
-->  2002  ACGACTCAGA      480
     2003
     2004  TCTTGTTGTC ACCCCCATTA TTGACAATCC AAAGGTGCAG AAAGCACTCT
-->  2005  GACAATTCCA      540
     2006
     2007  ATTGCTAATC CAGCATTTGT GGATAGCTGC AAACTGCGAT ATTGCTGATG
-->  2008  AGCGCTTTGA      600
     2009
     2010  CGCCACATTC CACACTAACG TGTTGGTGAA TTCTTCTGGG CATTGCCAGT
-->  2011  ACCTGCCTCC      660
     2012
     2013  AGGCATATTC AAGAGTTCCT GCTACATCG                        689
     2014
```

```
     2015  (2) INFORMATION FOR SEQ ID NO:103:
     2016
     2017      (i) SEQUENCE CHARACTERISTICS:
-->  2018          (A) LENGTH: 641 base pairs
     2019          (B) TYPE: nucleic acid
     2020          (C) STRANDEDNESS: single
     2021          (D) TOPOLOGY: linear
     2022
     2023      (ii) MOLECULE TYPE: other nucleic acid
     2024          (A) DESCRIPTION: /desc = "DNA"
     2025
     2026
     2027
     2028
     2029      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:
     2030
     2031  CAGGCCGCCA CATAGCTCCC GCCAAGTCCT CGGTGCCCCT TGCCATTTTC
-->  2032  CAGCCGCGCT       60
     2033
     2034  CCCACGAGGG TCACGGCGGC GGGGAGAGGT GGAGCCGCGA GAGCTCGGCC
-->  2035  GGGGGCCCCG      120
     2036
     2037  CCTGGTGGCC GCGGCCATGA CAGCGGCTCG GGACTGGCTC CTTTTCCGCG
-->  2038  CCCCTCCCGC      180
     2039
     2040  CGGAGGTGAG GGGAAGATGT CCATGTCAGG GTTCAAGGCC AAACCGAAGT
-->  2041  TACTGGCCTC      240
     2042
     2043  TATCTTCCAG GAGAACCAGG AGCCACAGCC GCGGCTCACG CCCCACCGCA
-->  2044  ACATTAAGAT      300
     2045
     2046  TACAAGTGGA CACCTGAGTC AGCAGGACCT GGAATCCCAG ATGAGAGAGC
-->  2047  TTATCTACAC      360
     2048
     2049  GACTCAGATC TTGTTGTCAC CCCCATTATT GACAATCCAA AGGTGCAGAA
-->  2050  AGCACTCTGA      420
     2051
     2052  CAAATAATGA AACAACCACC ATCGGTTAAA TTTGATGCAA AAATATTGCA
-->  2053  TCTACCAGCA      480
```

INPUT SET: S30866.raw

```
2054
2055    TTTTCAGTTC CAATTGCTAA TCCAGCATTT GTGGATAGCT GCAAACTGCG
-->  2056    ATATTGCTGA        540
2057
2058    TGAGCGCTTT GACGCCACAT TCCACACTAA CGTGTTGGTG AATTCTTCTG
-->  2059    GGCATTGCCA       ·600
2060
2061    ·GTACCTGCCT CCAGGCATAT TCAAGAGTTC CTGCTACATC G                        641
2062
```

```
2081    (2) INFORMATION FOR SEQ ID NO:105:·
2082
2083         (i) SEQUENCE CHARACTERISTICS:
-->  2084             (A) LENGTH: 140 base pairs
2085             (B) TYPE: nucleic acid
2086             (C) STRANDEDNESS: single
2087             (D) TOPOLOGY: linear
2088
2089        (ii) MOLECULE TYPE: other nucleic acid
2090             (A) DESCRIPTION: /desc = "DNA"
2091
2092
2093
2094
2095        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:
2096
2097    TGTCCNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->  2098    NNNNNNNNNN        60
2099
2100    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->  2101    NNNNNNNNNN       120
2102
2103    NNNNNNNNNN NNNNGACGTG                                                 140
2104
```

```
2123    (2) INFORMATION FOR SEQ ID NO:107:
2124
2125         (i) SEQUENCE CHARACTERISTICS:
-->  2126             (A) LENGTH: 44 base pairs
2127             (B) TYPE: nucleic acid
2128             (C) STRANDEDNESS: single
2129             (D) TOPOLOGY: linear
2130
2131        (ii) MOLECULE TYPE: other nucleic acid
2132             (A) DESCRIPTION: /desc = "DNA"
2133
2134
2135
2136
2137        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:
2138
2139    GATGAGNNNN NNNNNNNNNN· NNNNNNNNNN NNNNNNNCA AATG
2140        44
```

RAW SEQUENCE LISTING
PATENT APPLICATION *US/08/956,518*

INPUT SET: S30866.raw

```
        2141

        2160   (2) INFORMATION FOR SEQ ID NO:109:
        2161
        2162       (i) SEQUENCE CHARACTERISTICS:
-->     2163           (A) LENGTH: 110 base pairs
        2164           (B) TYPE: nucleic acid
        2165           (C) STRANDEDNESS: single
        2166           (D) TOPOLOGY: linear
        2167
        2168      (ii) MOLECULE TYPE: other nucleic acid
        2169           (A) DESCRIPTION: /desc = "DNA"
        2170
        2171
        2172
        2173
        2174      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:
        2175
        2176   TCTTGGNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->     2177   NNNNNNNNNN       60
        2178
        2179   NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNAACAG
        2180             110
        2181

        2200   (2) INFORMATION FOR SEQ ID NO:111:
        2201
        2202       (i) SEQUENCE CHARACTERISTICS:
-->     2203           (A) LENGTH: 80 base pairs
        2204           (B) TYPE: nucleic acid
        2205           (C) STRANDEDNESS: single
        2206           (D) TOPOLOGY: linear
        2207
        2208      (ii) MOLECULE TYPE: other nucleic acid
        2209           (A) DESCRIPTION: /desc = "DNA"
        2210
        2211
        2212
        2213
        2214      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:
        2215
        2216   GCTGATNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->     2217   NNNNNNNNNN       60
        2218
        2219   NNNNNNNNNN NNNCCTCCAG                                    80
        2220

        2239   (2) INFORMATION FOR SEQ ID NO:113:
        2240
        2241       (i) SEQUENCE CHARACTERISTICS:
-->     2242           (A) LENGTH: 168 base pairs
        2243           (B) TYPE: nucleic acid
        2244           (C) STRANDEDNESS: single
```

same

same

INPUT SET: S30866.raw

```
2245              (D) TOPOLOGY: linear
2246
2247         (ii) MOLECULE TYPE: other nucleic acid
2248              (A) DESCRIPTION: /desc = "DNA"
2249
2250
2251
2252
2253         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:
2254
2255    GCATANNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->2256    NNNNNNNNNN        60
2257
2258    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->2259    NNNNNNNNNN       120
2260
2261    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NCTAGTGG
2262         168
2263
```

```
2282    (2) INFORMATION FOR SEQ ID NO:115:
2283
2284         (i) SEQUENCE CHARACTERISTICS:
-->2285             (A) LENGTH: 195 base pairs
2286             (B) TYPE: nucleic acid
2287             (C) STRANDEDNESS: single
2288             (D) TOPOLOGY: linear
2289
2290        (ii) MOLECULE TYPE: other nucleic acid
2291             (A) DESCRIPTION: /desc = "DNA"
2292
2293
2294
2295
2296        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:
2297
2298    GAATCNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->2299    NNNNNNNNNN        60
2300
2301    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->2302    NNNNNNNNNN       120
2303
2304    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->2305    NNNNNNNNNN       180
2306
2307    NNNNNNNNTC CCTGG                                        195
2308
```

```
2327    (2) INFORMATION FOR SEQ ID NO:117:
2328
2329         (i) SEQUENCE CHARACTERISTICS:
-->2330             (A) LENGTH: 87 base pairs
2331             (B) TYPE: nucleic acid
```

INPUT SET: S30866.raw

```
      2332              (C) STRANDEDNESS: single
      2333              (D) TOPOLOGY: linear
      2334
      2335         (ii) MOLECULE TYPE: other nucleic acid
      2336              (A) DESCRIPTION: /desc = "DNA"
      2337
      2338
      2339
      2340
      2341         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:
      2342
      2343    GGATANNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->   2344    NNNNNNNNNN         60
      2345
      2346    NNNNNNNNNN NNNNNNNNNN TTGATAG                                    87
      2347
```

```
      2366    (2) INFORMATION FOR SEQ ID NO:119:
      2367
      2368         (i) SEQUENCE CHARACTERISTICS:
-->   2369              (A) LENGTH: 110 base pairs
      2370              (B) TYPE: nucleic acid
      2371              (C) STRANDEDNESS: single
      2372              (D) TOPOLOGY: linear
      2373
      2374         (ii) MOLECULE TYPE: other nucleic acid
      2375              (A) DESCRIPTION: /desc = "DNA"
      2376
      2377
      2378
      2379
      2380         (xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:
      2381
      2382    CCCAGNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->   2383    NNNNNNNNNN         60
      2384
      2385    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNAAGTGG
      2386              110
      2387
```

```
      2406    (2) INFORMATION FOR SEQ ID NO:121:
      2407
      2408         (i) SEQUENCE CHARACTERISTICS:
-->   2409              (A) LENGTH: 519 base pairs
      2410              (B) TYPE: nucleic acid
      2411              (C) STRANDEDNESS: single
      2412              (D) TOPOLOGY: linear
      2413
      2414         (ii) MOLECULE TYPE: other nucleic acid
      2415              (A) DESCRIPTION: /desc = "DNA"
      2416
      2417
      2418
```

RAW SEQUENCE LISTING
PATENT APPLICATION   *US/08/956,518*

```
        2419
        2420      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:
        2421
        2422    ACCAGANNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->     2423    NNNNNNNNNN      60
        2424
        2425    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->     2426    NNNNNNNNNN     120
        2427
        2428    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->     2429    NNNNNNNNNN     180
        2430
        2431    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->     2432    NNNNNNNNNN     240
        2433
        2434    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->     2435    NNNNNNNNNN     300
        2436
        2437    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->     2438    NNNNNNNNNN     360
        2439
        2440    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->     2441    NNNNNNNNNN     420
        2442
        2443    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
-->     2444    NNNNNNNNNN     480
        2445
        2446    NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNN
        2447    519
        2448
```

same

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/08/956,518*

INPUT SET: S30866.raw

Line	Error	Original Text
188	Entered (337) and Calc. Seq. Length (87) differ	(A) LENGTH: 337 base pairs
202	# of Sequences for line conflicts w/ running total	CTGCGAATGA 60
205	# of Sequences for line conflicts w/ running total	CGCTGCAGCC 120
208	# of Sequences for line conflicts w/ running total	AACCTGCTGT 180
211	# of Sequences for line conflicts w/ running total	TCTGGGGTAG 240
214	# of Sequences for line conflicts w/ running total	GGCGGGCAAC 300
239	Entered (54) and Calc. Seq. Length (0) differ	(A) LENGTH: 54 base pairs
924	Entered (48) and Calc. Seq. Length (0) differ	(A) LENGTH: 48 base pairs
943	Entered (41) and Calc. Seq. Length (0) differ	(A) LENGTH: 41 base pairs
1573	Entered (55) and Calc. Seq. Length (5) differ	(A) LENGTH: 55 base pairs
1587	# of Sequences for line conflicts w/ running total	GCACG 55
1754	Entered (457) and Calc. Seq. Length (107) differ	(A) LENGTH: 457 base pairs
1768	# of Sequences for line conflicts w/ running total	CAGAGGCGCG 60
1771	# of Sequences for line conflicts w/ running total	ACCCAGCGCC 120
1774	# of Sequences for line conflicts w/ running total	GAGACGCTGG 180
1777	# of Sequences for line conflicts w/ running total	TCACGTGGAG 240
1780	# of Sequences for line conflicts w/ running total	CGCGCGAGCC 300
1783	# of Sequences for line conflicts w/ running total	AGACGTGGAG 360
1786	# of Sequences for line conflicts w/ running total	GGCGTCTGGC 420
1793	Entered (308) and Calc. Seq. Length (58) differ	(A) LENGTH: 308 base pairs
1807	# of Sequences for line conflicts w/ running total	CAGCCGCGTC 60
1810	# of Sequences for line conflicts w/ running total	GGGGCCCCGC 120
1813	# of Sequences for line conflicts w/ running total	CCCTCCCGCC 180
1816	# of Sequences for line conflicts w/ running total	ACTGGCCCTC 240
1819	# of Sequences for line conflicts w/ running total	ACATTAAGGT 300
1826	Entered (145) and Calc. Seq. Length (45) differ	(A) LENGTH: 145 base pairs
1840	# of Sequences for line conflicts w/ running total	AGATGAGAGA 60
1843	# of Sequences for line conflicts w/ running total	AAAGGTGCAG 120
1850	Entered (84) and Calc. Seq. Length (34) differ	(A) LENGTH: 84 base pairs
1864	# of Sequences for line conflicts w/ running total	TATTGCATCT 60
1871	Entered (67) and Calc. Seq. Length (17) differ	(A) LENGTH: 67 base pairs
1885	# of Sequences for line conflicts w/ running total	GCGATATGTA 60
1892	Entered (100) and Calc. Seq. Length (50) differ	(A) LENGTH: 100 base pairs
1906	# of Sequences for line conflicts w/ running total	TTGGTGAATT 60
1931	Entered (392) and Calc. Seq. Length (92) differ	(A) LENGTH: 392 base pairs
1945	# of Sequences for line conflicts w/ running total	CAGAGGCGCG 60
1948	# of Sequences for line conflicts w/ running total	ACCCAGCGCC 120
1951	# of Sequences for line conflicts w/ running total	GAGACGCTGG 180
1954	# of Sequences for line conflicts w/ running total	TCACGTGGAG 240
1957	# of Sequences for line conflicts w/ running total	CGCGCGAGCC 300
1960	# of Sequences for line conflicts w/ running total	AGACGTGGAG 360
1967	Entered (689) and Calc. Seq. Length (139) differ	(A) LENGTH: 689 base pairs
1981	# of Sequences for line conflicts w/ running total	CTGCGGCTTG 60
1984	# of Sequences for line conflicts w/ running total	CGCAGGCCGC 120
1987	# of Sequences for line conflicts w/ running total	CTCCCACGAG 180
1990	# of Sequences for line conflicts w/ running total	CGCCTGGTGG 240
1993	# of Sequences for line conflicts w/ running total	GCCGGAGGTG 300
1996	# of Sequences for line conflicts w/ running total	TCTATCTTCC 360

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION *US/08/956,518*

INPUT SET: S30866.raw

Line	Error	Original Text	
1999	# of Sequences for line conflicts w/ running total	ATTACAAGTG	420
2002	# of Sequences for line conflicts w/ running total	ACGACTCAGA	480
2005	# of Sequences for line conflicts w/ running total	GACAATTCCA	540
2008	# of Sequences for line conflicts w/ running total	AGCGCTTTGA	600
2011	# of Sequences for line conflicts w/ running total	ACCTGCCTCC	660
2018	Entered (641) and Calc. Seq. Length (141) differ	(A) LENGTH: 641 base pairs	
2032	# of Sequences for line conflicts w/ running total	CAGCCGCGCT	60
2035	# of Sequences for line conflicts w/ running total	GGGGGCCCCG	120
2038	# of Sequences for line conflicts w/ running total	CCCCTCCCGC	180
2041	# of Sequences for line conflicts w/ running total	TACTGGCCTC	240
2044	# of Sequences for line conflicts w/ running total	ACATTAAGAT	300
2047	# of Sequences for line conflicts w/ running total	TTATCTACAC	360
2050	# of Sequences for line conflicts w/ running total	AGCACTCTGA	420
2053	# of Sequences for line conflicts w/ running total	TCTACCAGCA	480
2056	# of Sequences for line conflicts w/ running total	ATATTGCTGA	540
2059	# of Sequences for line conflicts w/ running total	GGCATTGCCA	600
2084	Entered (140) and Calc. Seq. Length (40) differ	(A) LENGTH: 140 base pairs	
2098	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	60
2101	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	120
2126	Entered (44) and Calc. Seq. Length (0) differ	(A) LENGTH: 44 base pairs	
2163	Entered (110) and Calc. Seq. Length (10) differ	(A) LENGTH: 110 base pairs	
2177	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	60
2203	Entered (80) and Calc. Seq. Length (30) differ	(A) LENGTH: 80 base pairs	
2217	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	60
2242	Entered (168) and Calc. Seq. Length (20) differ	(A) LENGTH: 168 base pairs	
2256	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	60
2259	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	120
2285	Entered (195) and Calc. Seq. Length (45) differ	(A) LENGTH: 195 base pairs	
2299	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	60
2302	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	120
2305	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	180
2330	Entered (87) and Calc. Seq. Length (37) differ	(A) LENGTH: 87 base pairs	
2344	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	60
2369	Entered (110) and Calc. Seq. Length (10) differ	(A) LENGTH: 110 base pairs	
2383	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	60
2409	Entered (519) and Calc. Seq. Length (80) differ	(A) LENGTH: 519 base pairs	
2423	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	60
2426	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	120
2429	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	180
2432	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	240
2435	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	300
2438	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	360
2441	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	420
2444	# of Sequences for line conflicts w/ running total	NNNNNNNNNN	480